A Provably Secure Composite Diffie–Hellman Scheme

M. I. Anokhin

Abstract. We suggest a modification of the Diffie–Hellman key distribution scheme with composite modulus and prove that this modification is secure under the general integer factoring intractability assumption. In this scheme, the modulus is the product of two arbitrary distinct primes. We show that if there exists a probabilistic polynomial-time algorithm that breaks the scheme with nonnegligible probability, then there also exists a probabilistic polynomial-time algorithm for factoring the modulus with nonnegligible probability. This improves the results obtained by Shmuely and McCurley [McC].

Keywords: key distribution, Diffie–Hellman scheme, integer factoring.

1. Introduction

The main problem in the theory of the Diffie–Hellman key distribution scheme [DH] is proving its security under some standard cryptographic assumption. For the original Diffie–Hellman scheme (with prime modulus), den Boer [B] and Maurer [Mau] proved the security under some assumptions apparently stronger than the general discrete logarithm intractability assumption. Note that Maurer's result is applicable not only to the original Diffie–Hellman scheme, but also to some its generalizations.

Shmuely [S] and McCurley [McC] considered variations of the *Composite Diffie–Hellman scheme*, i.e., the Diffie–Hellman scheme with the modulus n which is the product of two distinct primes. They proved (see [McC]) that any (probabilistic) algorithm $\mathfrak{A}$ breaking the scheme with nonzero probability can be used to construct a probabilistic algorithm $\mathfrak{B}$ for factoring the modulus n with probability at least $1/2$. However, Shmuely's result (in the form cited in [McC]) does not allow us to state that if $\mathfrak{A}$ has polynomial running time and nonnegligible probability of success, then $\mathfrak{B}$ also runs in polynomial time. This is only true if, e.g., the prime factors of the modulus n have a certain special form. McCurley's result [McC] guarantees polynomial running time for $\mathfrak{B}$ provided that $\mathfrak{A}$ runs in polynomial time and has nonnegligible probability of success. But in the McCurley scheme, the prime factors of the modulus n have to be chosen in some special way. This means that the security of this scheme is based on an apparently stronger assumption than the general factoring intractability assumption, namely, on the intractability assumption for factoring products of two distinct primes chosen in this way.

This paper suggests another modification of the Composite Diffie–Hellman key distribution scheme. The modulus n in our scheme is the product of two *arbitrary* distinct primes. The base in this scheme is taken uniformly at random from the set of all elements of odd order in $\mathbb{Z}_n^*$. Both the modulus and the base are chosen by a trusted authority. Our main result is that if there exists a probabilistic algorithm $\mathfrak{A}$ that breaks our scheme with probability at least $\varepsilon(n)$, then there also exists a probabilistic algorithm $\mathfrak{B}$ for factoring the modulus with probability at least $\varkappa\varepsilon(n)$, where $\varkappa$ is an absolute positive constant. Moreover, if $\mathfrak{A}$ runs in polynomial time, then $\mathfrak{B}$ also runs in *polynomial* time. The probability of success of $\mathfrak{B}$ can be increased (in particular, to $1/2$, as in the results of Shmuely and McCurley) by a standard iterating procedure. As a corollary, we prove the security of the suggested scheme under the general factoring intractability assumption. To our knowledge, this is the first proving the security of the Diffie–Hellman scheme under some standard cryptographic assumption. The proof of our result develops the proofs of results of Shmuely and McCurley [McC].

It should be noted that the randomness of the base is essential for the proof of our result, whereas in the McCurley scheme, the base is fixed. Note also that in our scheme, all algorithms (including initializing performed by the trusted authority) are efficient.

In Sec. 2, we give the notation used in this paper. Section 3 describes our scheme. In Sec. 4, we state the main results of the paper. Section 5 contains some auxiliary lemmas which are used in the proof of the main theorem. Finally, in Sec. 6, we define some probability distributions and prove the main theorem.

2. Notation

In this paper, we use the following notation:

- log denotes the logarithm to the base 2;

- $[x]$ denotes the integral part of a real number x;
- $\mathrm{pln}(x)$ denotes some function of the form cx^d, where c, d are real positive constants (in different places, pln may denote different functions);
- $\mathbb{N}$ denotes the set $\{1, 2, \dots\}$ of all positive integers;
- $\langle a \rangle$ denotes the (cyclic) group generated by the element a;
- $\operatorname{ord} a$ denotes the order of the element a of a group;
- $\delta(n, a)$ denotes $[n/\operatorname{ord} a]$, where $n \in \mathbb{N}$ and a is an element of finite order of a group;
- $\mathbb{Z}_n$ denotes the set $\{0, 1, \dots, n-1\}$ usually considered as a ring with respect to addition and multiplication modulo n ($n \in \mathbb{N}$);
- $\mathbb{Z}_n^*$ denotes the set of all numbers in $\mathbb{Z}_n$ which are coprime to n, i.e., the group of all units in the ring $\mathbb{Z}_n$;
- H_n denotes the set (subgroup) of all elements of odd order in $\mathbb{Z}_n^*$;
- $1^{(l)}$ denotes the string of l binary ones.

The notation $e \in_{\mathrm{R}} E$ means that e is an uniformly distributed random element of the finite set E. We assume that these random elements are mutually independent.

All our results are stated for the uniform model of computation, i.e., by the word "algorithm", we mean a probabilistic Turing machine.

3. The Scheme

Let us describe our key distribution scheme. Denote the security parameter as l ($l \in \mathbb{N}$). We assume that there exists a polynomial-time algorithm $\mathfrak{G}$ that, given $1^{(l)}$, outputs a two-element set $\{p, q\}$ of primes such that $\log p, \log q \geq \mathrm{pln}(l)$. Our scheme requires a trusted authority which is responsible for maintaining a certified public directory.

To initialize the scheme, the trusted authority runs the following algorithm:

(1) generate $\{p, q\} = \mathfrak{G}\left(1^{(l)}\right)$ and compute the *modulus* $n = pq$;
(2) compute α and β such that $p - 1 = 2^\alpha p'$ and $q - 1 = 2^\beta q'$, where p' and q' are odd;
(3) choose $a \in_{\mathrm{R}} \{1, \dots, p-1\}$, $b \in_{\mathrm{R}} \{1, \dots, q-1\}$ and compute (unique) $g \in \mathbb{Z}_n^*$ (the *base*) such that $g \equiv a^{2^\alpha} \pmod{p}$ and $g \equiv b^{2^\beta} \pmod{q}$ (it is easily seen that $a^{2^\alpha} \bmod p \in_{\mathrm{R}} H_p$, $b^{2^\beta} \bmod q \in_{\mathrm{R}} H_q$ and hence, $g \in_{\mathrm{R}} H_n$);
(4) make n and g public by placing them in the certified public directory (all other results of random choices and computations must be secret).

Any user (say, A) starts with taking n and g from the certified public directory, selecting $x_{\mathrm{A}} \in_{\mathrm{R}} \mathbb{Z}_n$, computing $y_{\mathrm{A}} = g^{x_{\mathrm{A}}} \bmod n$, and sending it to the trusted authority, keeping x_{A} secret. Then the trusted authority forms a pair (name of A, y_{A}) and puts it in the certified public directory.

To generate a common secret key, two users A and B compute $g^{x_{\mathrm{A}} x_{\mathrm{B}}} \bmod n$ as $y_{\mathrm{B}}^{x_{\mathrm{A}}} \bmod n$ and $y_{\mathrm{A}}^{x_{\mathrm{B}}} \bmod n$, respectively, where y_{B} (resp., y_{A}) is taken from the certified public directory.

It is evident that all the above algorithms run in time polynomial in l.

Remark. Let n be the modulus in the scheme. It is easy to see that $z \to z^{2^{[\log m]}} \bmod m$ is a homomorphism of $\mathbb{Z}_m^*$ onto H_m for any $m \in \mathbb{N}$. Therefore, if $z \in_{\mathrm{R}} \mathbb{Z}_n$, then $z^{2^{[\log n]}} \bmod n \in_{\mathrm{R}} H_n$ provided that $z \in \mathbb{Z}_n^*$, which is true with probability $|\mathbb{Z}_n^*|/|\mathbb{Z}_n| \geq 1/4$ and can be verified in polynomial time (e.g., by the Euclidean algorithm). Hence, there exists a polynomial-time algorithm that, given only n, outputs $g' \in_{\mathrm{R}} H_n$ with probability at least $1/4$. Moreover, we can distinguish between successful an unsuccessful runs of this algorithm. The probability of success can be increased to 1 by a standard iterating procedure, but then the algorithm will run only in expected polynomial time. Thus, given only the modulus n, we can efficiently generate an element having the same distribution as the base g in the described scheme. This means that the knowledge of g does not help in factoring n.

4. Main Results

In this section, we formulate the main results of this paper. In the following main theorem, n denotes the product of two arbitrary distinct primes p and q. Let also S, T, and ε denote some real-valued functions defined on the set of such n's, and $\varkappa$ denotes an absolute positive constant (in particular, we may take $\varkappa = 8/225$).

Theorem. *Let $\mathfrak{A}$ be an algorithm that, given n, g, $g^{x_1} \bmod n$, and $g^{x_2} \bmod n$, where $g \in_R H_n$, $x_1 \in_R \mathbb{Z}_n$, $x_2 \in_R \mathbb{Z}_n$, outputs a list of size at most $S(n)$ such that*

$$\Pr\{g^{x_1 x_2} \bmod n \in \mathfrak{A}(n, g, g^{x_1} \bmod n, g^{x_2} \bmod n)\} \geq \varepsilon(n),$$

where the probability is taken over g, x_1, x_2, and random bits used by $\mathfrak{A}$. Assume that the running time of the algorithm $\mathfrak{A}$ is at most $T(n)$ on any input (n, g, g_1, g_2), where $g, g_1, g_2 \in \mathbb{Z}_n$. Then there exists an algorithm $\mathfrak{B}$ that, given n, outputs the set $\{p, q\}$ with probability at least $\varkappa\varepsilon(n)$, where the probability is taken over random bits used by $\mathfrak{B}$. The running time of the algorithm $\mathfrak{B}$ on any input n is at most $\operatorname{pln}(\log n)(T(n) + S(n))$.

The most interesting case of this theorem is when $T(n) \leq \operatorname{pln}(\log n)$ and $\varepsilon(n) \geq 1/\operatorname{pln}(\log n)$. Then, obviously, $S(n) \leq \operatorname{pln}(\log n)$. Hence, if the algorithm $\mathfrak{A}$ outputs a list of size at most $S(n)$ containing the desired element $g^{x_1 x_2} \bmod n$ with probability at least $\varepsilon(n)$, then we can find this element with probability at least $\varepsilon(n)/S(n) \geq 1/\operatorname{pln}(\log n)$ by choosing an element of the list uniformly at random. Therefore, in the corollary stated below, we assume that the algorithm $\mathfrak{A}$ outputs the common secret key rather than the list. Moreover, in this case, we may assume that a breaking algorithm (in the definition stated below) runs in polynomial time only on a valid public information of the scheme, whereas in the theorem, we have to suppose that the running time of the algorithm $\mathfrak{A}$ is at most $T(n)$ on any input (n, g, g_1, g_2) with arbitrary $g, g_1, g_2 \in \mathbb{Z}_n$. Indeed, if $c[\log n]^d$, where $c, d \in \mathbb{N}$, is an upper bound for the running time of a breaking algorithm on a valid public information of the scheme, then we can modify this algorithm in such a way that it counts the made steps and aborts the run when the running time exceeds $c[\log n]^d$.

Definition. The above key distribution scheme is *insecure* if there exists a polynomial-time algorithm that, given only the public information of the scheme (i.e., n, g, $g^{x_A} \bmod n$, $g^{x_B} \bmod n$), outputs the common secret key $g^{x_A x_B} \bmod n$ with probability at least $1/\operatorname{pln}(l)$ for infinitely many l's. Otherwise the above scheme is *secure*.

Corollary. *Suppose that the above key distribution scheme is insecure. Then*

(1) *for any real-valued function ρ defined on $\mathbb{N}$ and such that $\rho(k) \leq 1 - 2^{-\operatorname{pln}(k)}$ for each $k \in \mathbb{N}$, there exists a polynomial-time algorithm $\mathfrak{B}'$ that, given $n = pq$, where $\{p, q\} = \mathfrak{G}\left(1^{(l)}\right)$, outputs the set $\{p, q\}$ with probability at least $\rho(l)$ for infinitely many l's;*

(2) *there exists an algorithm $\mathfrak{B}''$ that, given $n = pq$, where $\{p, q\} = \mathfrak{G}\left(1^{(l)}\right)$, outputs the set $\{p, q\}$ with probability 1 in expected polynomial running time for infinitely many l's.*

In the definition and the corollary, the probabilities are taken over the joint distribution of input data and random bits used by the algorithms.

Informally speaking, the corollary means that if there are no efficient algorithms for factoring $n = pq$, where $\{p, q\} = \mathfrak{G}\left(1^{(l)}\right)$, then the above key distribution scheme is secure.

The corollary can be easily derived from the theorem by using a standard iterating procedure.

Let $\mathfrak{A}$ be an algorithm satisfying the theorem hypotheses. Then the factoring algorithm $\mathfrak{B}$ is as follows (n denotes an input number pq, where p and q are distinct primes):

Algorithm $\mathfrak{B}$

1. If n is even, then output the set $\mathfrak{B}(n) = \{2, n/2\}$ and successfully terminate. Otherwise, choose $g_0 \in_R \mathbb{Z}_n$, $e_1 \in_R \{1, 3, \ldots, n-2\}$, $e_2 \in_R \{1, 3, \ldots, n-2\}$.
2. For $i = 0, 1, \ldots, [\log n] - 1$, do stages 2.1–2.3:
   2.1. Put $g_i = g_0^{2^i} \bmod n$.
   2.2. Run $\mathfrak{A}$ on input $\left(n, g_i^4 \bmod n, g_i^{2e_1} \bmod n, g_i^{2e_2} \bmod n\right)$. If the algorithm $\mathfrak{A}$ requests the kth random bit which was not used in its previous calls, then this bit is stored after its generation for using as the kth random bit not only in the current call, but in all the following ones as well. In other words, in all its calls, $\mathfrak{A}$ uses the same random binary string to which the new random bits are appended as needed. Denote the output list of the algorithm $\mathfrak{A}$ on the above input as $\{a_{i1}, a_{i2}, \ldots, a_{is_i}\}$ (this list may be empty).
   2.3. For every $j \in \{1, 2, \ldots, s_i\}$ such that $a_{ij} \in \mathbb{Z}_n$, compute $p_{ij} = \gcd(n, a_{ij} - (g_i^{e_1 e_2} \bmod n))$.
3. If we have $p_{ij} \neq 1$ and $p_{ij} \neq n$ for some $i \in \mathbb{Z}_{[\log n]}$, $j \in \{1, 2, \ldots, s_i\}$, then output the set $\mathfrak{B}(n) = \{p_{ij}, n/p_{ij}\}$ which evidently coincides with $\{p, q\}$. In this case, the run of $\mathfrak{B}(n)$ is successful; otherwise it is unsuccessful.

It is obvious that the running time of $\mathfrak{B}$ on any input n of the above form is at most $\mathrm{pln}(\log n) \times (T(n) + S(n))$.

5. Auxiliary Lemmas

In this section, we assume that $n = pq$, where p and q are arbitrary distinct odd primes such that $p - 1 = 2^\alpha p'$ and $q - 1 = 2^\beta q'$, where p' and q' are odd. Then $\mathbb{Z}_n^*$ is a direct product $U \times V \times H$, where $U = \langle u \rangle$ and $V = \langle v \rangle$ are cyclic subgroups of orders 2^α and 2^β, respectively, and $H = H_n$. Moreover, $u^{2^{\alpha-1}} \bmod n$, $v^{2^{\beta-1}} \bmod n$, and $u^{2^{\alpha-1}} v^{2^{\beta-1}} \bmod n$ are all elements of order 2 in $\mathbb{Z}_n^*$. We assume u and v to be chosen such that $u^{2^{\alpha-1}} v^{2^{\beta-1}} \equiv -1 \pmod n$.

Let $\varphi\colon \{x \in \mathbb{Z}_n^* \mid \operatorname{ord} x \text{ is even}\} \to \mathbb{Z}_n^*$ denote the function defined as follows. Let $x \in \mathbb{Z}_n^*$, $\operatorname{ord} x = 2^\gamma d$, where $\gamma \geq 1$ and d is odd. Then $x\varphi = x^{2^{\gamma-1}} \bmod n$. Thus, the order of $x\varphi$ modulo H is exactly 2.

For an arbitrary $h \in H$, let $X(h)$ denote the set of all $x \in \mathbb{Z}_n^*$ such that

(1) $\operatorname{ord} x$ is even;
(2) $(x\varphi)^4 \bmod n = h$;
(3) $-1 \bmod n \notin \langle x \rangle$.

Let also $\mu = \min\{\alpha, \beta\}$.

Lemma 1. *For any $h \in H$,*

$$|X(h)| = 2^{\alpha+\beta} - \frac{4^\mu + 2}{3}.$$

Proof. Fix $h \in H$. Define a function $\psi\colon U \times V \to \mathbb{Z}_n^*$ as follows: $y\psi = yh_1 \bmod n$, where y is an element of $U \times V$ and h_1 is the element of H such that $h_1^{2 \operatorname{ord} y} \bmod n = h$ (h_1 exists and is unique because $|H|$ is odd).

Let us show that if $Y = \{ab \bmod n \mid a \in U,\ b \in V,\ \operatorname{ord} a \neq \operatorname{ord} b\}$, then $Y\psi = X(h)$. Let $y = ab \bmod n$ (where $a \in U$, $b \in V$) be an element of Y. It is evident that $\operatorname{ord}(y\psi) = \operatorname{ord} y \operatorname{ord} h_1$, where $\operatorname{ord} y$ is a nonzero power of 2 (if $\operatorname{ord} y = 1$, then $\operatorname{ord} a = \operatorname{ord} b = 1$) and $\operatorname{ord} h_1$ is odd. Therefore, $\operatorname{ord}(y\psi)$ is even and $(y\psi\varphi)^4 \bmod n = (yh_1)^{2 \operatorname{ord} y} \bmod n = h_1^{2 \operatorname{ord} y} \bmod n = h$. Now, suppose that $(y\psi)^k \equiv -1 \pmod n$ for some $k = 2^\xi k'$, where k' is odd. Then we have $a^k \equiv u^{2^{\alpha-1}} \pmod n$ and $b^k \equiv v^{2^{\beta-1}} \pmod n$. This means that $\operatorname{ord} a \mid 2k$ and $\operatorname{ord} b \mid 2k$, but $\operatorname{ord} a \nmid k$ and $\operatorname{ord} b \nmid k$. Hence, $\operatorname{ord} a = 2^{\xi+1} = \operatorname{ord} b$ (since $\operatorname{ord} a$ and $\operatorname{ord} b$ are powers of 2), which contradicts the definition of Y. Thus, $y\psi \in X(h)$, and we have proved that $Y\psi \subseteq X(h)$.

Let now $x = abh_2 \bmod n$ (where $a \in U$, $b \in V$, $h_2 \in H$) be an element of $X(h)$. We show that $y = ab \bmod n \in Y$ and $y\psi = x$. Indeed, $\operatorname{ord} x = \operatorname{ord} y \operatorname{ord} h_2$, where $\operatorname{ord} y$ is a nonzero power of 2 (since $\operatorname{ord} x$ is even) and $\operatorname{ord} h_2$ is odd. If $\operatorname{ord} a = \operatorname{ord} b = 2^\eta$, then $\eta \geq 1$ (since $\operatorname{ord} y = \max\{\operatorname{ord} a, \operatorname{ord} b\} = 2^\eta$) and

$$x^{2^{\eta-1} \operatorname{ord} h_2} \equiv \left(a^{2^{\eta-1}} b^{2^{\eta-1}}\right)^{\operatorname{ord} h_2} \equiv \left(u^{2^{\alpha-1}} v^{2^{\beta-1}}\right)^{\operatorname{ord} h_2} \equiv (-1)^{\operatorname{ord} h_2} \equiv -1 \pmod n$$

(here we use the observation that $u^{2^{\alpha-1}} \bmod n$ and $v^{2^{\beta-1}} \bmod n$ are the only elements of order 2 in the subgroups U and V, respectively). This contradicts the definition of $X(h)$, hence, we have $\operatorname{ord} a \neq \operatorname{ord} b$, i.e., $y \in Y$. It is obvious that $h_2^{2 \operatorname{ord} y} \bmod n = x^{2 \operatorname{ord} y} \bmod n = \left(x^{(\operatorname{ord} y)/2}\right)^4 \bmod n = (x\varphi)^4 \bmod n = h$, therefore, $y\psi = x$. Thus, $X(h) \subseteq Y\psi$, and we have $Y\psi = X(h)$. It is easily seen that the function ψ is one-to-one (because the product UVH is direct), therefore, the latter equality implies $|Y| = |X(h)|$.

Let us compute $|Y|$. It is easy to see that if C_{2^ν} is a cyclic group of order 2^ν, then

$$\left|\left\{z \in C_{2^\nu} \mid \operatorname{ord} z = 2^i\right\}\right| = \begin{cases} 1 & \text{if } i = 0; \\ 2^{i-1} & \text{if } 1 \leq i \leq \nu. \end{cases}$$

Therefore, we have

$$\begin{aligned} |X(h)| = |Y| &= |U \times V| - |\{(a,b) \in U \times V \mid \operatorname{ord} a = \operatorname{ord} b\}| \\ &= |U||V| - \sum_{i=0}^{\mu} \left|\left\{a \in U \mid \operatorname{ord} a = 2^i\right\}\right| \left|\left\{b \in V \mid \operatorname{ord} b = 2^i\right\}\right| \\ &= 2^{\alpha+\beta} - 1^2 - \sum_{i=1}^{\mu} \left(2^{i-1}\right)^2 = 2^{\alpha+\beta} - \frac{4^\mu + 2}{3}. \end{aligned}$$

□

Lemma 2. *For any $h \in H$,*

$$\frac{|X(h)||H|}{n} \geq \frac{2}{9}.$$

Proof. First we substitute the values of $|X(h)|$ (from Lemma 1), n, and $|H|$ into the left-hand side of the desired inequality:

$$\frac{|X(h)||H|}{n} = \frac{(2^{\alpha+\beta} - (4^\mu + 2)/3)p'q'}{n} = \frac{\left(2^{\alpha+\beta} - (4^\mu + 2)/3\right) p'q'}{2^{\alpha+\beta}p'q'} \times \frac{(p-1)(q-1)}{pq} = \left(1 - \frac{4^\mu + 2}{3 \cdot 2^{\alpha+\beta}}\right)\left(1 - \frac{1}{p}\right)\left(1 - \frac{1}{q}\right). \quad (1)$$

The inequalities $\alpha \geq \mu$ and $\beta \geq \mu$ imply $\alpha + \beta \geq 2\mu$ and $2^{\alpha+\beta} \geq 4^\mu$. Therefore, we have

$$1 - \frac{4^\mu + 2}{3 \cdot 2^{\alpha+\beta}} \geq 1 - \frac{4^\mu + 2}{3 \cdot 4^\mu}, \quad (2)$$

where $\mu \geq 1$ and hence, $4^\mu \geq 4$. Since $f(t) = 1 - (t+2)/(3t)$ is an increasing function of t for $t > 0$, we have

$$1 - \frac{4^\mu + 2}{3 \cdot 4^\mu} = f(4^\mu) \geq f(4) = \frac{1}{2}. \quad (3)$$

The inequalities $p \geq 3$ and $q \geq 3$ imply

$$1 - \frac{1}{p} \geq \frac{2}{3} \quad \text{and} \quad 1 - \frac{1}{q} \geq \frac{2}{3}. \quad (4)$$

Now the lemma follows from (1)–(4). □

Lemma 3. *For any $h \in H$,*

$$\frac{[\delta(n,h)/2]}{((n-1)/2)((\delta(n,h)+1)/n)} \geq \frac{2}{5}.$$

Proof. For brevity, put $\delta = \delta(n,h)$. It is evident that $[t] > t - 1$ for each real t. Therefore, we have

$$\frac{[\delta/2]}{((n-1)/2)((\delta+1)/n)} > \frac{(\delta/2) - 1}{((n-1)/2)((\delta+1)/n)} = \frac{n}{n-1}\frac{\delta - 2}{\delta + 1}. \quad (5)$$

Note that $\operatorname{ord} h$ divides $|H| = p'q'$, hence, we have

$$\operatorname{ord} h \leq p'q' = \frac{(p-1)(q-1)}{2^{\alpha+\beta}} < \frac{n}{4}.$$

Thus, $\delta = [n/\operatorname{ord} h] \geq 4$. It is easily seen that $F(t) = (t-2)/(t+1)$ is an increasing function of t for $t > -1$, hence, $(\delta - 2)/(\delta + 1) = F(\delta) \geq F(4) = 2/5$. Moreover, it is obvious that $n/(n-1) > 1$. To prove the lemma, it suffices to substitute these estimates in (5). □

6. Proof of the Main Theorem

Let us define some probability distributions. If a is an element of finite order of a group and $n \in \mathbb{N}$, then we denote the distribution (over the group $\langle a \rangle$) of a^i, where $i \in_{\mathrm{R}} \mathbb{Z}_n$, as $D_{a,n}$. It is easy to see that if $k = n \bmod \operatorname{ord} a$, then we have

$$\Pr_{D_{a,n}}\{a^i\} = \begin{cases} (\delta(n,a)+1)/n & \text{if } 0 \leq i \leq k-1; \\ \delta(n,a)/n & \text{if } k \leq i \leq (\operatorname{ord} a) - 1. \end{cases} \quad (6)$$

Similarly to the uniform distribution, the notation $b \in_{D_{a,n}} \langle a \rangle$ means that b is a random element of the group $\langle a \rangle$ with respect to the distribution $D_{a,n}$. We assume that each such a random element is selected independently of others.

For $n \in \mathbb{N}$, denote the probability space of all triples (h, h_1, h_2), where $h \in_{\mathrm{R}} H_n$, $h_1 \in_{D_{h,n}} \langle h \rangle$, $h_2 \in_{D_{h,n}} \langle h \rangle$, as Ω_n. For an arbitrary $(h, h_1, h_2) \in \Omega_n$, we have

$$\Pr\{(h, h_1, h_2)\} = \Pr\{h\} \Pr_{D_{h,n}}\{h_1\} \Pr_{D_{h,n}}\{h_2\}, \tag{7}$$

where $\Pr\{h\} = 1/|H_n|$.

Similarly to [McC], let $DH(n, h, h_1, h_2) = h^{x_1 x_2} \bmod n$, where $n \in \mathbb{N}$, $h \in \mathbb{Z}_n^*$, $h_1 = h^{x_1} \bmod n \in \langle h \rangle$, and $h_2 = h^{x_2} \bmod n \in \langle h \rangle$. It is evident that DH is a well-defined function. The task of an adversary in the above key distribution scheme (see Sec. 3) is to compute $DH(n, \omega)$, given $n = pq$, where $\{p, q\} = \mathfrak{G}\left(1^{(l)}\right)$, and a random element ω of the space Ω_n.

Let $\mathfrak{A}$ be an algorithm satisfying the theorem hypotheses and $\mathfrak{B}$ the algorithm constructed from $\mathfrak{A}$ as in Sec. 4. Since the probability of success of $\mathfrak{B}$ on any even valid input n is $1 \geq \varkappa\varepsilon(n)$, here and below we assume that n is the product of two distinct odd primes p and q. For convenience, we sometimes write random binary strings used by algorithms as a part of their input data. For a given n, we may assume that the algorithm $\mathfrak{A}$ uses random string $r \in_{\mathrm{R}} \{0,1\}^{T(n)}$ and the algorithm $\mathfrak{B}$ uses random parameters $g_0 \in_{\mathrm{R}} \mathbb{Z}_n$, $e_1 \in_{\mathrm{R}} \{1, 3, \ldots, n-2\}$, $e_2 \in_{\mathrm{R}} \{1, 3, \ldots, n-2\}$, and $r \in_{\mathrm{R}} \{0,1\}^{T(n)}$ (recall that $\mathfrak{A}$ uses the same random string in all its calls from $\mathfrak{B}$).

Let M_n be the set of all $(\omega, r) \in \Omega_n \times \{0,1\}^{T(n)}$ such that the run of $\mathfrak{A}$ on the input (n, ω) using the random string r is successful, i.e.,

$$M_n = \left\{ (\omega, r) \in \Omega_n \times \{0,1\}^{T(n)} \;\middle|\; DH(n, \omega) \in \mathfrak{A}(n, \omega, r) \right\}.$$

Lemma 4. *The run of $\mathfrak{B}(n, g_0, e_1, e_2, r)$ is successful if $g_0 \in \mathbb{Z}_n^*$, $\operatorname{ord} g_0$ is even, $-1 \bmod n \notin \langle g_0 \rangle$, and $\left((g_0\varphi)^4 \bmod n, (g_0\varphi)^{2e_1} \bmod n, (g_0\varphi)^{2e_2} \bmod n, r\right) \in M_n$ (see notation in Sec. 5).*

Proof. For brevity, denote $g_0\varphi$ by f. Then we have $f = g_i$ for some $i \in \mathbb{Z}_{[\log n]}$, hence, the list $\{a_{i1}, a_{i2}, \ldots, a_{is_i}\}$ computed at stage 2.2 of the algorithm $\mathfrak{B}$ contains

$$a_{ij} = DH\left(n, f^4 \bmod n, f^{2e_1} \bmod n, f^{2e_2} \bmod n\right) = f^{4m^2 e_1 e_2} \bmod n,$$

where m is a number such that $f^{4m} \equiv f^2 \pmod{n}$ (recall that $f^2 \in H_n$ (see Sec. 5), therefore, $\left\langle f^4 \bmod n \right\rangle = \left\langle f^2 \bmod n \right\rangle$). Note that $f^{8m^2} \equiv \left(f^{4m}\right)^{2m} \equiv \left(f^2\right)^{2m} \equiv f^{4m} \equiv f^2 \pmod{n}$, hence, we have

$$a_{ij}^2 \equiv f^{8m^2 e_1 e_2} \equiv f^{2e_1 e_2} \equiv (f^{e_1 e_2})^2 \pmod{n}.$$

This means that

$$(a_{ij} - (f^{e_1 e_2} \bmod n))(a_{ij} + (f^{e_1 e_2} \bmod n)) \equiv 0 \pmod{n}. \tag{8}$$

Suppose that $a_{ij} - (f^{e_1 e_2} \bmod n)$ is divisible neither by p nor q. In this case, $a_{ij} - (f^{e_1 e_2} \bmod n)$ is coprime to n, and (8) implies $a_{ij} + (f^{e_1 e_2} \bmod n) \equiv 0 \pmod{n}$. But then $-1 \bmod n = a_{ij}^{-1} f^{e_1 e_2} \bmod n \in \langle f \rangle \subseteq \langle g_0 \rangle$, which contradicts the hypothesis. If $a_{ij} - (f^{e_1 e_2} \bmod n)$ is divisible both by p and q, then we have $f^{4m^2 e_1 e_2} \bmod n = a_{ij} = f^{e_1 e_2} \bmod n$. But $f^{4m^2 e_1 e_2} \bmod n \in H_n$ and $f^{e_1 e_2} \bmod n \notin H_n$ since f has order 2 modulo H_n (see Sec. 5) and $e_1 e_2$ is odd. Therefore, $a_{ij} - (f^{e_1 e_2} \bmod n)$ is divisible by exactly one prime factor of n. This implies $p_{ij} \neq 1$ and $p_{ij} \neq n$, i.e., the run of $\mathfrak{B}(n, g_0, e_1, e_2, r)$ is successful. □

Proof of the Theorem. From the theorem hypotheses and (7), we obtain

$$\begin{aligned}
\varepsilon(n) \leq \Pr M_n &= \sum_{(h,h_1,h_2,r) \in M_n} \Pr\{(h, h_1, h_2, r)\} \\
&= \sum_{(h,h_1,h_2,r) \in M_n} \Pr\{(h, h_1, h_2)\} \Pr\{r\} \\
&= \sum_{(h,h_1,h_2,r) \in M_n} \frac{1}{|H_n|} \Pr_{D_{h,n}}\{h_1\} \Pr_{D_{h,n}}\{h_2\} \frac{1}{2^{T(n)}} \\
&= \frac{1}{|H_n|\, 2^{T(n)}} \sum_{(h,h_1,h_2,r) \in M_n} \Pr_{D_{h,n}}\{h_1\} \Pr_{D_{h,n}}\{h_2\}.
\end{aligned} \tag{9}$$

For an arbitrary $m = (h, h_1, h_2, r) \in M_n$, let us define the set Z_m as the set of all quadruples (g_0, e_1, e_2, r) such that

(1) $g_0 \in X(h)$;
(2) $e_1, e_2 \in \{1, 3, \ldots, n-2\}$;
(3) $(h')^{e_1} \equiv h_1 \pmod n$, $(h')^{e_2} \equiv h_2 \pmod n$, where $h' \in H_n$ is such that $(h')^2 \equiv h \pmod n$ (h' exists and is unique because $|H_n|$ is odd).

Let a be an element of odd order of some group and $b = a^k \in \langle a \rangle$, where $k \in \mathbb{Z}_{\operatorname{ord} a}$. Then the parities of numbers in the sequence $k, k + \operatorname{ord} a, k + 2 \operatorname{ord} a, \ldots, k + (\delta(n, a) - 1) \operatorname{ord} a$, in which all elements are distinct and belong to $\{t \in \mathbb{Z}_n \mid a^t = b\}$, alternate. Hence, we have

$$\left|\left\{t \in \{1, 3, \ldots n-2\} \mid a^t = b\right\}\right| \geq \left[\frac{\delta(n, a)}{2}\right].$$

The latter inequality implies that for any $m = (h, h_1, h_2, r) \in M_n$,

$$|Z_m| \geq |X(h)| \left[\frac{\delta(n, h')}{2}\right]^2 = |X(h)| \left[\frac{\delta(n, h)}{2}\right]^2,$$

where $h' \in H_n$ is such that $(h')^2 \equiv h \pmod n$, as above. Note that for each $m = (h, h_1, h_2, r) \in M_n$ and each $(g_0, e_1, e_2, r) \in Z_m$, $\operatorname{ord} g_0$ is even, $-1 \bmod n \notin \langle g_0 \rangle$ (since $g_0 \in X(h)$) and

$$\begin{aligned} h &= (g_0 \varphi)^4 \bmod n, \\ h_1 &= (g_0 \varphi)^{2e_1} \bmod n, \\ h_2 &= (g_0 \varphi)^{2e_2} \bmod n \end{aligned}$$

(since $h' = (g_0 \varphi)^2 \bmod n$). Therefore, by Lemma 4, the run of $\mathfrak{B}(n, g_0, e_1, e_2, r)$ is successful. Moreover, $Z_{m_1} \cap Z_{m_2} = \emptyset$ for $m_1 \neq m_2$ ($m_1, m_2 \in M_n$) because m can be recovered uniquely from an arbitrary element of Z_m.

Let us estimate the probability of success of the run of $\mathfrak{B}(n, g_0, e_1, e_2, r)$:

$$\begin{aligned} &\Pr\{\text{the run of } \mathfrak{B}(n, g_0, e_1, e_2, r) \text{ is successful}\} \geq \frac{\left|\bigcup\limits_{m \in M_n} Z_m\right|}{n((n-1)/2)^2 2^{T(n)}} \\ &= \frac{1}{n((n-1)/2)^2 2^{T(n)}} \sum_{(h, h_1, h_2, r) \in M_n} |Z_{(h, h_1, h_2, r)}| \\ &\geq \frac{1}{n((n-1)/2)^2 2^{T(n)}} \sum_{(h, h_1, h_2, r) \in M_n} |X(h)| \left[\frac{\delta(n, h)}{2}\right]^2 \\ &= \frac{1}{|H_n|\, 2^{T(n)}} \sum_{(h, h_1, h_2, r) \in M_n} \frac{|X(h)||H_n|}{n} \frac{[\delta(n, h)/2]^2}{((n-1)/2)^2 \Pr_{D_{h,n}}\{h_1\} \Pr_{D_{h,n}}\{h_2\}} \\ &\quad \times \Pr_{D_{h,n}}\{h_1\} \Pr_{D_{h,n}}\{h_2\} \quad (\text{note that } \Pr_{D_{h,n}}\{h_1\}, \Pr_{D_{h,n}}\{h_2\} \geq \delta(n, h)/n > 0 \text{ by (6)}) \\ &\geq \frac{1}{|H_n|\, 2^{T(n)}} \sum_{(h, h_1, h_2, r) \in M_n} \left(\frac{|X(h)||H_n|}{n}\right) \left(\frac{[\delta(n, h)/2]}{((n-1)/2)((\delta(n, h) + 1)/n)}\right)^2 \\ &\quad \times \Pr_{D_{h,n}}\{h_1\} \Pr_{D_{h,n}}\{h_2\} \quad (\text{since } \Pr_{D_{h,n}}\{h_1\}, \Pr_{D_{h,n}}\{h_2\} \leq (\delta(n, h) + 1)/n \text{ by (6)}) \\ &\geq \frac{8}{225} \frac{1}{|H_n|\, 2^{T(n)}} \sum_{(h, h_1, h_2, r) \in M_n} \Pr_{D_{h,n}}\{h_1\} \Pr_{D_{h,n}}\{h_2\} \quad (\text{by Lemmas 2, 3}) \\ &\geq \frac{8}{225} \varepsilon(n) \quad (\text{by (9)}). \quad \square \end{aligned}$$

7. Acknowledgement

The author thanks N. P. Varnovsky for attentive reading early versions of this paper and many helpful discussions and comments.

References

[B] B. den Boer, "Diffie–Hellman is as strong as discrete log for certain primes", Proc. CRYPTO'88, *Lect. Notes in Comp. Sci.*, **403**, 530–539 (1989).

[DH] W. Diffie and M. E. Hellman, "New directions in cryptography", *IEEE Trans. on Inform. Theory*, **22**, No. 6, 644–654 (1976).

[Mau] U. M. Maurer, "Towards the equivalence of breaking the Diffie–Hellman protocol and computing discrete logarithms", Proc. CRYPTO'94, *Lect. Notes in Comp. Sci.*, **839**, 271–281.

[McC] K. S. McCurley, "A key distribution system equivalent to factoring", *J. of Cryptology*, **1**, No. 2, 95–105 (1988).

[S] Z. Shmuely, "Composite Diffie–Hellman public-key generating systems are hard to break", Tech. Report No. 356, Comp. Sci. Dept., Technion-Israel Institute of Technology, Feb. 1985.